Evaluation of alkylating pyrrole-imidazole polyamide conjugates by a novel method for high-throughput sequencer

AUTHOR(S):

Kashiwazaki, Gengo; Maeda, Rina; Kawase, Takashi; Hashiya, Kaori; Bando, Toshikazu; Sugiyama, Hiroshi

CITATION:

Kashiwazaki, Gengo ...[et al]. Evaluation of alkylating pyrrole-imidazole polyamide conjugates by a novel method for high-throughput sequencer. Bioorganic and Medicinal Chemistry 2018, 26(1): 1-7

ISSUE DATE:

2018-01-01

URL:

http://hdl.handle.net/2433/230869

Evaluation of alkylating pyrrole-imidazole polyamide conjugates by a novel method for high-throughput sequencer

Gengo Kashiwazaki[a], Rina Maeda[b], Takashi Kawase[c], Kaori Hashiya[a], Toshikazu Bando[a,*], Hiroshi Sugiyama[a,d,*]

[a]Department of Chemistry, Graduate School of Science, Kyoto University, Kitashirakawa-oiwake-cho, Sakyo, Kyoto 606-8502, Japan
[b]Graduate School of Advanced Integrated Studies in Human Survivability, Kyoto University, Nakaadachi-cho, Yoshida, Sakyo, Kyoto 606-8306, Japan
[c]Department of Systems Science, Graduate School of Informatics, Kyoto University, Yoshida-Honmachi 36-1, Sakyo, Kyoto, 606-8501, Japan
[d]Institute for Integrated Cell-Material Sciences (iCeMS), Kyoto University, Yoshida-ushinomiya-cho, Sakyo, Kyoto 606-8501, Japan
*Corresponding authors. Tel.: +81 75 753 4002; fax +81 75 753 3670.
E-mail address: hs@kuchem.kyoto-u.ac.jp (H. Sugiyama)
bando@kuchem.kyoto-u.ac.jp (T. Bando)

ABSTRACT

N-methylpyrrole-*N*-methylimidazole (PI) polyamides are a class of DNA minor groove binders with DNA sequence-specificity. DNA-alkylating PI polyamide conjugates are attractive candidates as anticancer drugs acting through DNA damage and its subsequent inhibition of cell proliferation. One example is a chlorambucil-PI polyamide conjugate targeting the runt-related transcription factor (RUNX) family. RUNX1 has pro-oncogenic properties in acute myeloid leukemia, and recently the chlorambucil-PI polyamide conjugate was demonstrated to have anticancer effects. Herein, we apply another DNA-alkylating agent, *seco*-CBI, to target the consensus sequence of the RUNX family. Two types of CBI conjugates were prepared and their binding properties were characterized by Bind-n-Seq analysis using a high-throughput sequencer. The sequencing data were analyzed by two methods, MERMADE and our new MR (motif identification with a reference sequence), and the resultant binding motif logos were as predicted from the pairing rules proposed by Dervan et al. This is the first report to employ the MR method on alkylating PI polyamide conjugates. Moreover, cytotoxicity of conjugates **3** and **4** against a human non-small cell lung cancer, A549, were examined to show promising IC_{50}s of 120 nM and 63 nM, respectively. These findings suggest *seco*-CBI-PI polyamide conjugates are candidates for oncological therapy.

1. Introduction

There are a number of DNA-alkylating agents applied to cancer therapy,[1,2] one of which is chlorambucil, a nitrogen mustard. The alkylation chemistry targets N7 of guanine and N3 of adenine, and interstrand crosslinking that specifically occurs between guanines of 5′-GCC-3′/3′-CGG-5′.[3] Although chlorambucil has been long used as a drug for chronic lymphatic leukemia and Hodgkin's lymphoma, there are known adverse effects, such as nausea, vomiting, myelosuppression, diarrhea, and tremors, because specific organ toxicities derive from the high proliferative rates of bone-marrow cells and the epithelium of the gastrointestinal tract. Therefore, much effort has been put into alleviating these side effects. One of the strategies is to confer DNA sequence-specificity to the functional molecule. It can be assumed that delivery of an effector to a predetermined target would produce the desired biological results while reducing the off-target effects, even if the carrier cannot restrict its destination to one site in the billions of base-pairs. *N*-methylpyrrole-*N*-methylimidazole (PI) polyamide is a well-characterized synthetic molecule that binds in the minor groove and recognizes the DNA sequence following the pairing rule.[4-7] PI polyamide with chlorambucil on its γ-turn was shown to alkylate and interstrand crosslink DNA in a sequence-specific manner in vitro and caused cell cycle arrest in the G_2/M phase.[8] The chlorambucil conjugate was later shown to regulate expression of histone H4 mRNA in various cancer cell lines and to produce a loss of tumorigenicity in mouse xenograft models of cancer.[9] Recently PI polyamide-chlorambucil conjugate Chb-M′ was reported to show a remarkable improvement of overall survival periods in mouse models of acute myeloid leukemia compared with a DMSO control, chlorambucil, cytarabine (Cylocide), and PI polyamide control.[10] Cytarabine is an anticancer drug for chemotherapy of acute and chronic myelogenous leukemia, acute lymphocytic leukemia, acute promyelocytic leukemia, and Hodgkin's lymphoma. The target of Chb-M′ is the RUNX family, which is strongly associated with cancer development.[11] Chb-M′ exerts its antitumor effects through inhibiting the binding of RUNX family members to the promoter regions of the p53 degrading enzymes *BCL11A* and *TRIM24*, thereby inducing apoptosis.

Inspired by these unprecedented results, we synthesized and evaluated *seco*-CBI-PI polyamide conjugates. *seco*-CBI (1-chloromethyl-5-hydroxy-1,2-dihydro-3*H*-benz[*e*]-indole)[12,13] is a synthetic DNA alkylator based on Duocarmycin A, and its conjugate with PI polyamide has long been investigated.[14,15] While chlorambucil has a flexible alkyl linker, *seco*-CBI is conjugated with PI polyamide via a rigid indole linker, and its alkylation site would be defined more specifically. Another characteristic of *seco*-CBI

conjugate is its greater cytotoxicity compared with the chlorambucil conjugate.[16] In 2015, mutant KRAS suppression by *seco*-CBI conjugate resulted in tumor death in an oncogenic mutation-specific manner.[17-19] These findings corroborate the significance of *seco*-CBI conjugate as a promising drug candidate.

High-throughput sequencing is widely used for analysis of binding motifs of PI polyamides.[20] One of the tested *seco*-CBI-PI polyamide conjugates is shown in Figure 1.[21] The Bind-n-Seq[22] substrate DNA comprises 102 base pairs containing two adapters, a barcode, and 21 base-pair random sequences.[18,19,21,23-25] The workflow starts from primer extension to prepare dsDNA containing random sequences, then a binding/alkylation reaction by the PI polyamide, affinity purification by exploiting the interaction between biotinylated PI polyamide and streptavidin magnetic beads, followed by emulsion PCR and sequencing. These properties are seen in vitro, but experiments targeting genomic DNA in vivo, Chem-Seq[26], enabled biological discussions at gene levels.[21,27,28] Here we investigated binding preferences of two PI polyamide conjugates using Bind-n-Seq analysis with a new MR (motif identification with a reference sequence) for the first time for alkylating conjugates as well as with the conventional MERMADE. In addition, their cytotoxicity against an A549 cell line was explored to expand the options for cancer treatment to targeting the RUNX family.

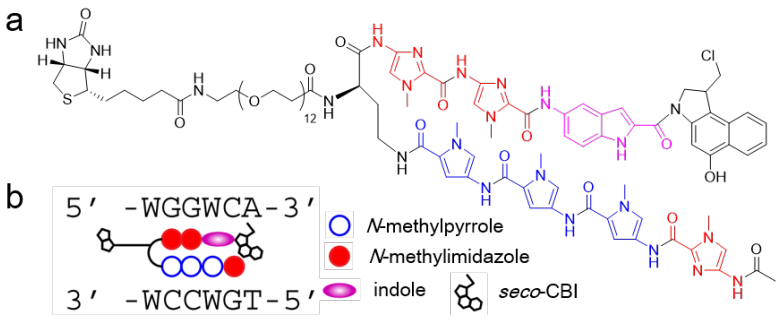


Figure 1. A chemical structure (a) and schematic representation (b) of a biotinylated *seco*-CBI-PI polyamide conjugate.[21]

2. Materials and methods

2.1. General

Reagents and solvents were purchased from standard suppliers and used without further purification. Automated polyamide synthesis was performed on a PSSM-8 system (Shimadzu) on 30 μmol scale by Fmoc chemistry.[29,30] HPLC purification was performed with a JASCO PU-2080 Plus pump, a UV-2075 Plus detector (254 nm), an MX-2080-32 mixer and a DG-2080-54 degasser. A Chemcobond 5-ODS-H column (4.6 x 150 mm;

Chemco Plus Scientific, Osaka, Japan) was used; the mobile phase was a gradient of acetonitrile with TFA (0.1%, v/v in water) at a flowrate of 1.0 mL min^{-1}. ESI-TOFMS data were obtained on a BioTOF II (Bruker Daltonics). MALDI-TOFMS was performed on a microflex-KSII (Bruker Daltonics). NMR Spectra were recorded on a JEOL JNM ECA-600 NMR spectrometer. The following abbreviations apply to spin multiplicity: s (singlet), d (doublet), t (triplet), m (multiplet).

2.2. Synthesis

2.2.1. Fmoc-Py-CO_2H loading to 2-Cl-Trt resin

Fmoc-Py-CO_2H was loaded to 2-chlorotrityl chloride resin (2-Cl-Trt-Cl) resin mainly as previously reported,[31] but *N*-methylpyrrolidone was chosen as a solvent during the coupling reaction instead of DMF for better solubility.

2.2.2. Compound 5

53 mg of 2-Cl-Trt-Cl resin (37.5 µmol/100 mg) loading AcIPII-(*R*)-^{BocHN}γ-PP was shaken in a solution of 150 µL of hexafluoroisopropanol and 350 µL of CH_2Cl_2 for 3 h at room temperature. Then the solution was dripped into Et_2O to obtain the crude solid (17.8 mg). HPLC retention time: 13.6 min (0-100% over 20 min). MALDI-TOFMS *m/z* calculated for $C_{44}H_{54}N_{17}O_{11}^+$ $[M+H]^+$ 996.418, found 996.271. ^{1}H NMR (600 MHz, DMSO-d_6): δ = 10.43 (s, 1H), 10.26 (s, 1H), 9.93 (s, 1H), 9.89 (s, 1H), 9.86 (s, 1H), 9.40 (s, 1H), 8.21 (t, J = 5.8 Hz, 1H), 7.63 (s, 1H), 7.52 (s, 1H), 7.42 (s, 1H), 7.40 (m, 2H), 7.18 (d, J = 2.0 Hz, 1H), 7.15 (apparent s, 1H), 7.02 (d, J = 7.5 Hz, 1H), 6.91 (d, J = 1.4 Hz, 1H), 6.82 (d, J = 2.0 Hz, 1H), 4.10 (m, 1H), 4.01 (s, 3H), 3.95 (s, 3H), 3.94 (s, 3H), 3.87 (s, 3H), 3.82 (s, 3H), 3.80 (s, 3H), 3.29 (m, 2H), 2.02 (s, 3H), 1.91 (m, 1H), 1.79 (m, 1H), 1.39 (s, 9H).

2.2.3. Compound 6

0.86 mg (0.86 µmol) of **5** and 0.45 mg (2.0 eq, 1.7 µmol) of PyBOP were dissolved in 8.6 µL of DMF, and to that solution was added 1.8 µL (12 eq, 10 µmol) of DIEA. After vortexing, 0.44 mg (1.0 eq, 0.86 µmol) of aminoindole-*seco*-CBI (TFA salt)[32] was added and shaken for 2 h at room temperature. Et_2O was added to workup and the obtained solid was proceeded to the next reaction without purification. HPLC retention time: 14.0 min (0-100% over 20 min).

2.2.4. Compound 7: 250 µL of 50% TFA in CH_2Cl_2 was added to the compound **6**. 5 min later, Et_2O was added to obtain the crude solid (1.43 mg). HPLC retention time: 11.4 min (0-100% over 20 min). MALDI-TOFMS m/z calculated for $C_{61}H_{61}ClN_{20}O_{10}^+$ $[M + H]^+$

1291.446, found 1291.482.

2.2.5. Conjugate 1

11 µL (1.1 mg, 1.1 µmol) of DMF solution of NHS-PEG_{12}-Biotin, 1.0 µL (6.0 µmol) of DIEA and 8 µL of DMF were added to 1.43 mg of the crude compound **6** and shaken for 30 min at room temperature. The product solution was diluted by DMF and injected into HPLC for purification to yield 1.88 mg of the product powder. HPLC retention time: 12.0 min (0-100% over 20 min), 10.3 min (30-90% over 20 min). MALDI-TOFMS m/z calculated for $C_{98}H_{128}ClN_{23}NaO_{25}S^+$ $[M + Na]^+$ 2116.875, found 2116.876.

2.2.6. Compound 8

80 mg of 2-chlorotrityl chloride resin (2-Cl-Trt-Cl) resin (37.5 µmol/100 mg) loading AcIPII-(*R*)-^{BocHN}γ-PPP was shaken in a solution of 300 µL of hexafluoroisopropanol and 700 µL of CH_2Cl_2 for 3 h at room temperature. Then the solution was dripped into Et_2O to obtain the crude solid (28.3 mg).HPLC retention time: 13.4 min (0-100% over 20 min). MALDI-TOFMS m/z calculated for $C_{50}H_{60}N_{19}O_{12}^+$ $[M + H]^+$ 1118.466, found 1118.357. ^{1}H NMR (600 MHz, DMSO-d_6): δ = 10.43 (s, 1H), 10.26 (s, 1H), 9.94 (s, 1H), 9.91 (s, 1H), 9.884 (s, 1H), 9.877 (s, 1H), 9.40 (s, 1H), 8.22 (t, J = 5.5 Hz, 1H), 7.63 (s, 1H), 7.53 (s, 1H), 7.42 (s, 1H), 7.41 (d, J = 2.0 Hz, 1H), 7.39 (d, J = 1.3 Hz, 1H), 7.22 (d, J = 2.1 Hz, 1H), 7.18 (d, J = 1.4 Hz, 1H), 7.16 (apparent s, 1H), 7.04 (d, J = 1.4 Hz, 1H), 7.02 (d, J = 7.6 Hz, 1H), 6.92 (d, J = 2.1 Hz, 1H), 6.84 (d, J = 2.0 Hz, 1H), 4.11 (m, 1H), 4.01 (s, 3H), 3.95 (s, 3H), 3.94 (s, 3H), 3.87 (s, 3H), 3.834 (s, 3H), 3.831 (s, 3H), 3.82 (s, 3H), 3.29 (m, 2H), 2.02 (s, 3H), 1.92 (m, 1H), 1.79 (m, 1H), 1.40 (s, 9H).

2.2.7. Compound 9

1.22 mg (1.09 µmol) of **8** and 1.1 mg (2.0 eq, 2.2 µmol) of PyBOP were dissolved in 10 µL of DMF, and to that solution was added 2.1 µL (11 eq, 12 µmol) of DIEA. 10 min later, 0.55 mg (1.0 eq, 1.1 µmol) of aminoindole-*seco*-CBI (TFA salt)[32] was added and shaken for 5 h at room temperature. Et_2O was added to workup and though the property of the product was oily, it was proceeded to the next reaction without purification. HPLC retention time: 14.8 min (0-100% over 20 min). MALDI-TOFMS m/z calculated for $C_{72}H_{75}ClN_{22}NaO_{13}^+$ $[M + Na]^+$ 1513.547, found 1513.561.

2.2.8. Compound 10

320 µL of 50% TFA in CH_2Cl_2 was added to the compound **9**. 30 min later, Et_2O was added to obtain the dark brown powder (1.30 mg). HPLC retention time: 12.9 min (0-

100% over 20 min). MALDI-TOFMS m/z calculated for $C_{67}H_{68}ClN_{22}O_{11}^+$ [*M* + H]$^+$ 1391.512, found 1391.522.

2.2.9. Conjugate 2

8.9 µL (0.89 mg, 0.95 µmol) of DMF solution of NHS-PEG_{12}-Biotin, 0.89 µL (5.2 µmol) of DIEA and 8.9 µL of DMF were added to 1.30 mg of the crude compound **10** and shaken for 1 h at room temperature. The product solution was diluted by DMF and injected into HPLC for purification. This product was purified again to yield 0.40 mg of the product powder. HPLC retention time: 12.0 min (0-100% over 20 min), 11.0 min (30-90% over 20 min). MALDI-TOFMS m/z calculated for $C_{98}H_{128}ClN_{23}NaO_{25}S^+$ [*M* + Na]$^+$ 2238.923, found 2238.973.

2.2.10. Compound 11

AcIPII-γ-PP Trt resin was synthesized in a stepwise reaction by the Fmoc solid-phase protocol. Subsequently, the resin was treated with CH_2Cl_2 (350 µL) and 30% hexafluoroisopropanol (150 µL) at rt for 3 h. After filtration, the filtrate was evaporated and CH_2Cl_2 (400 µL), MeOH (100 µL) and Et_2O (5 mL) were added. The solution was sonicated and the target product was collected. A weight of the target product was 16.9 mg (19 µmol). MALDI-TOFMS *m/z* calcd for $C_{39}H_{45}N_{16}O_9^+$ [*M* + H]$^+$ 881.355, found 881.371. ^{1}H NMR (600 MHz, DMSO-d_6): δ=10.43 (s, 1H), 10.27 (s, 1H), 9.88 (s, 1H), 9.82 (s, 2H), 9.38 (s, 1H), 8.29 (t, J= 5.8 Hz, 1H), 7.63 (s, 1H), 7.52 (s, 1H), 7.43 (s, 1H), 7.40 (d, J= 1.4 Hz, 1H), 7.39 (s, 1H), 7.18 (d, J= 2.1 Hz, 1H), 7.14 (d, J= 1.4 Hz, 1H), 6.85 (apparent s, 1H), 6.81 (s, 1H), 4.01 (s, 3H), 3.96 (s, 3H), 3.95 (s, 3H), 3.87 (s, 3H), 3.81 (s, 6H), 3.27 (m, 2H), 2.28 (t, J= 7.6 Hz, 2H), 2.02 (s, 3H), 1.82 (m, 2H).

2.2.11. Conjugate 3

Compound **11** (4.0 mg, 4.6 µmol, 1.0 eq.), PyBOP (4.8 mg, 9.2 µmol, 2.0 eq.) and DIEA (4.8 µL, 28 µmol, 6.0 eq.) in DMF (10 µL) were stirred for 5 min. Then H_2N-indole–*seco*-CBI (3.6 mg, 9.2 µmol, 2.0 eq.) in DMF (5 µL) was added and stirred at rt for 1 h. After consumption of activated ester was confirmed by HPLC analysis. Et_2O was added to the mixture and resultant was collected by centrifuged, and washed by Et_2O and CH_2Cl_2. A weight of the target product was 2.8 mg (2.2 µmol, 49% yield) after HPLC purification. The retention time was 11.1 min (0.1% TFA containing 36–56% acetonitrile over a linear gradient for 30 min at a flow rate of 1.0 mL/min detected at 254 nm). ESI-TOFMS *m/z* calcd for $C_{61}H_{62}ClN_{19}O_{10}^{2+}$ [*M*+ 2H]$^{2+}$ 627.7302, found 627.7289.

2.2.12. Compound 12

AcIPII-γ-PPP Trt resin was synthesized in a stepwise reaction by the Fmoc solid-phase protocol. Subsequently, the resin was treated with CH_2Cl_2 (700 µL) and 30% hexafluoroisopropanol (300 µL) at rt for 3 h. After filtration, the filtrate was evaporated and CH_2Cl_2 (800 µL), MeOH (200 µL) and Et_2O (5 mL) were added. The solution was sonicated and the target product was collected. A weight of the target product was 29.0 mg (29 µmol). MALDI-TOFMS *m/z* calcd for $C_{45}H_{51}N_{18}O_{10}^+$ $[M + H]^+$ 1003.403, found 1003.422. ^{1}H NMR (600 MHz, DMSO-d_6): δ =10.43 (s, 1H), 10.27 (s, 1H), 9.89 (s, 1H), 9.88 (s, 2H), 9.83 (s, 1H), 9.39 (s,1H), 8.29 (t, J = 6.2 Hz, 1H), 7.63 (s, 1H), 7.52 (s, 1H), 7.43 (s, 1H), 7.42 (d, J = 1.4 Hz, 1H), 7.40 (d, J = 2.1 Hz, 1H), 7.22 (d, J = 1.4 Hz, 1H), 7.18 (d, J = 1.3 Hz, 1H), 7.16 (d, J = 2.0 Hz, 1H),7.04 (d, J = 2.0 Hz, 1H), 6.86 (d, J = 1.4 Hz, 1H), 6.84 (d, J = 2.0 Hz, 1H), 4.01 (s, 3H), 3.96 (s, 3H), 3.95 (s, 3H), 3.87 (s, 3H), 3.83 (s, 3H), 3.824 (s, 3H), 3.817 (s, 3H), 3.27 (m, 2H), 2.28 (t, J = 7.5 Hz, 2H), 2.02 (s, 3H), 1.82 (m, 2H).

2.2.13. Conjugate 4

Compound **12** (3.4 mg, 3.4 µmol, 1.0 eq.), PyBOP (3.5 mg, 6.8 µmol, 2.0 eq.) and DIEA (3.6 µL, 20 µmol, 6.0 eq.) in DMF (10 µL) were stirred for 5 min. Then H_2N-indole–*seco*-CBI (1.6 mg, 4.1 µmol, 1.2 eq.) in DMF (3 µL) was added and stirred at rt for 1 h. After consumption of activated ester was confirmed by HPLC analysis. Et_2O was added to the mixture and resultant was collected by centrifuged, and washed by Et_2O and CH_2Cl_2. A weight of the target product was 0.5 mg (0.36 µmol, 21% yield) after HPLC purification. The retention time was 15.7 min (0.1% TFA containing 30–70% acetonitrile over a linear gradient for 30 min at a flow rate of 1.0 mL/min detected at 254 nm). MALDI-TOFMS *m/z* calcd for $C_{67}H_{67}ClN_{21}O_{11}^+$ $[M + H]^+$ 1376.501, found 1376.505.

2.3. Bind-n-Seq assay

2.3.1. Primer extension

The ODN1 (5'-CCATCTCATCCCTGCGTGTCTCCGACTCAG CTAAGGTAAC N_{21} ATCACCGACTGCCCATAGAGAGGAAAGCGGAGGCGTAGTGG-3'), ODN2 (5'-CCATCTCATCCCTGCGTGTCTCCGACTCAG TAAGGAGAAC N_{21} ATCACCGACTGCCCATAGAGAGGAAAGCGGAGGCGTAGTGG-3') and ODN3 (5'-CCATCTCATCCCTGCGTGTCTCCGACTCAG AAGAGGATTC N_{21} ATCACCGACTGCCCATAGAGAGGAAAGCGGAGGCGTAGTGG-3') were purchased from Sigma-Aldrich. Ion Torrent adapter A1 is 5'-

CCATCTCATCCCTGCGTGTCTCCGACTCAG-3', P1 adapter is 5'-ATCACCGACTGCCCATAGAGAGGAAAGCGGAGGCGTAGTGG-3'. Ion Xpress Barcodes are 5'-CTAAGGTAAC-3', 5'-TAAGGAGAAC-3' and 5'-AAGAGGATTC-3', for ODN1, 2 and 3, respectively. N_{21} is the random sequence to be analyzed which is expected to contain the binding sites of the test PIP conjugate. Primer extensions for preparation of duplex of the ODNs were performed in a 25 µL reaction volume and consisting of 5.25 µL nuclease-free water, 12.5 µL GoTaq Green Master Mix, 4.5 µL of 50 µM primer (5'-CCACTACGCCTCCGCTTTCCTCTCTA-3'), 1.25 µL of 10 mM $MgCl_2$ and 1.5 µL of 50 µM template ODN 1, 2 or 3. The thermal condition was programmed for 2 min at 95 °C, 1 min, followed by 1 min at 63 °C, 4 min at 72 °C in ProFlex PCR System (Thermo Fisher Scientific). dsODNs1 and 2 were used as the substrates for conjugate **1** and **2**, respectively. dsODN3 was prepared as the input control during "Amplification" step described later, and proceeded to "Affinity purification".

2.3.2. Determination of PIP concentration

Molar extinction coefficients ε of the PIP conjugate in DMF solution at around 305 nm were calculated as 7.64×10^4 and 8.63×10^4, respectively[33] and their concentrations were determined by the Lambert-Beer law. PIP conjugate was dissolved in DMSO and ε in DMSO was assumed to be same as in DMF. Absorbance was measured by NanoDrop (ND-1000, Thermo Fisher Scientific). 1.0 µM of DMSO solution of PIP conjugate was diluted to 100 nM by TKMC solution (10 mM Tris-HCl pH 7.0, 10 mM KCl, 10 mM $MgCl_2$ and 5 mM $CaCl_2$).

2.3.3. Binding reaction

25 µL of that 100 nM solution was added to 25 µL of the primer extended solution, and the resulting solution was incubated for 16 h at rt.

2.3.4. Enrichment reactions

The procedure was based on a previous paper.[25] After the supernatant was removed, Streptavidin M-280 Dynabeads (50 µL, 0.5 mg/sample) were washed with 50 µL/sample of BSA solution (3.5 mg/mL in PBS) for 10 min twice, then with 50 µL/sample with calf thymus DNA solution (0.5 mg/mL in water) for 90 min. The supernatant was discarded and the beads were treated with 100 µL/sample of binding and washing buffer (5 mM Tris-HCl pH 7.5, 0.5 mM EDTA and 1 M NaCl) for 1 min thrice. Resuspension of 50 µL/sample of the beads in the binding and washing buffer was divided up among samples. The sample solutions from binding reaction were added to the buffer and incubated at

room temperature for 1 h with pipetting every 10 min. The supernatant was removed and the beads were washed with 0.5 mL of binding and washing buffer for 5 min, then with 0.5 mL of TKMC solution for 10 min twice. Enriched DNA was recovered by heating in 100 µL of elution buffer (2% SDS, 100 mM $NaHCO_3$ and 3 mM biotin) at 300 rpm at 65 ^oC for 7 h followed by magnetic separation.

2.3.5. Amplification

A PCR reaction was performed in a 25 µL volume using Ion Plus Fragment Library Kit (Thermo Fisher Scientific) following the manufacturer's instructions but the cycle number was set at 15.

2.3.6. Affinity purification

PCR products were purified by Agencourt AMPure XP beads (Beckman Coulter) following the manufacturer's instructions. The quality and concentration were analyzed by Agilent 2100 Bioanalyzer and High Sensitivity DNA Kit (Agilent Technologies). Due to a high concentration of the input control, it was diluted by a factor of 10 before loaded on the chip.

2.3.7. Emulsion PCR

Emulsion PCR of all the three sample was performed at a time on the Ion OneTouch 2 System (Thermo Fisher Scientific) using the Ion PGM Hi-Q View OT2 Kit. The library concentrations from both dsODN1 and 2 were 8 pM based on the results from Bioanalyzer. As that from dsODN3 showed a relatively broad peak, its 20 pM was mixed with the other two. The resulting beads were enriched by Ion OneTouch 2 ES.

2.3.8. Sequencing

The enriched libraries on Ion 318 Chip v2 BC were sequenced by Ion PGM System (ThermoFisher Scientific) using Ion PGM Hi-Q View Sequencing Kit.

2.3.9. Data analysis

Torrent Suite software 5.2.2 was applied to obtain fastq files and the following analyses were done by "Terminal" software in Ubuntu. The raw data were reanalyzed to fit into our purpose by setting the minimum read length as 8, then converted to fasta files (script: awk 'NR % 4 == 1 {print ">" $0} NR % 4 == 2 {print $0}'). These fasta files were trimmed to pick up only sequences with 21 bases and to remove reads which contained three continuous bases such as AAA (PRINSEQ script[34]: perl prinseq-lite.pl -min_len 21

-max_len 21 -custom_params "A 3;T 3;G 3;C 3" -fasta). PRINSEQ is available at https://sourceforge.net/projects/prinseq/files/standalone/. A particular number of random reads were subsampled so that the numbers of both the sample and input control would be the same (script: perl fasta-subsample <sequences> <count>). "fasta-subsample" is one of the other utilities in the MEME Suite available at http://meme-suite.org/doc/download.html?man_type=web. 7-mers were identified from sequences of the sample and control (MERMADE script: perl kmer_counter.pl -k 7). Enrichment values were acquired by comparison of the sample and control data (MERMADE script: perl kmer_selector.pl -d -r 1.5). The default threshold of enrichment was 2, but as the resulting sequences were too few to proceed to motif analysis, it was lowered to 1.5. Mermade matrices were generated (MERMADE scripts: perl mermade.pl), and visualized as logos by enoLOGOS[35] (http://www.benoslab.pitt.edu/cgi-bin/enologos/enologos.cgi). MERMADE scripts are available from the Korf lab (http://korflab.ucdavis.edu/Datasets/BindNSeq/).

2.3.10. Updated MR method for Bind-n-Seq analysis

The MR method reported previously[24] was slightly modified in order to detect sliding of sequences from their reference more accurately. To detail, sliding was judged by counting the match bases between the reference sequence and the sample: the reference sequence is fixed, the sample sequence was overlapped base by base from left end to right end, and the match bases are counted at each disposition. Its complementary sequence is examined in the same way at the same time. The location to achieve the maximum number of match bases is identified as the correct disposition. If that location overlaps completely with the reference sequence, the read is considered to be fit, and otherwise slid. In the original MR method, a sequenced read was judged as fit if either a sample or its complementary sequence overlaps completely, but this updated MR method finds out the best location from the sample and the complementary together. Thus, if the slid-state complementary sequence scores higher than the fit-state original one, that read will be judged as slid and excluded from the list to compose the resultant motif. This modification should attain more reasonable representation. This program and its source code are available from H. Sugiyama.

2.4. Cytotoxicity Assay

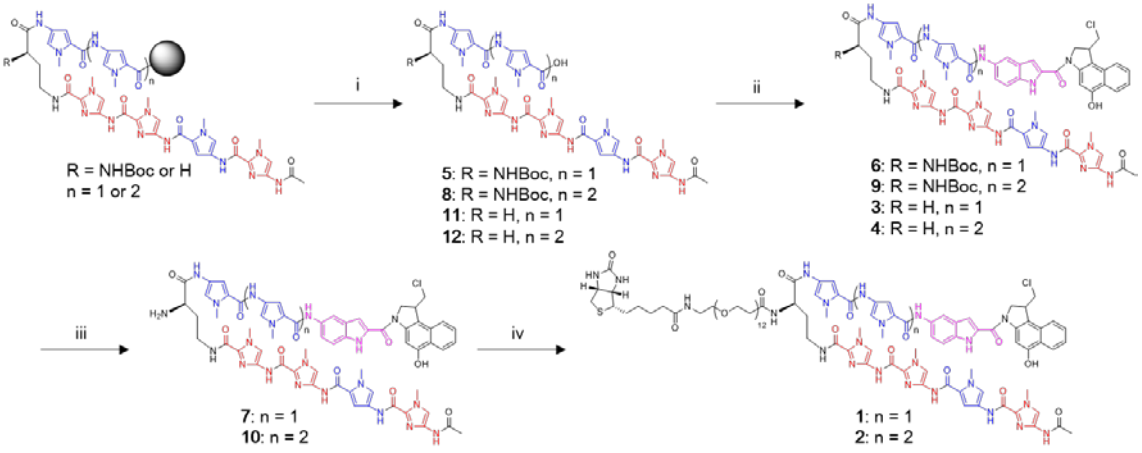
Human lung adenocarcinoma epithelial A549 cells (JCRB0076, establisher: Giard, D.) were obtained from JCRB cell bank. A549 was cultured in Dulbecco's Modified Eagle Medium (DMEM; Thermo Fisher) with 10% of FBS and 1% of L-glutamine. The cells

were maintained at 37 °C in humidified atmosphere of 95% air and 5% CO_2. Colorimetric assays using cell count reagent SF containing WST-8 (Nacalai tesque) were performed on 96-well plates. A 100 µL amount of cell suspension was added to each well (5000 cells/well) and incubated at 37 °C for 24 h. *seco*-CBI-PI polyamide conjugates were dissolved in DMSO, and 100 µL of each solution in the medium (final DMSO concentration was 0.2%) was added. After treatment for 48 h, removed the medium which contained conjugates and 100 µL of mixture of cell count reagent SF and medium (SF:medium = 1:10) was added to each well and incubated at 37 °C. Absorbance was then measured at 450 nm using a SpectraMax M2e microplate reader (Molecular Devices) after 1 h treatment.

3. Results and Discussion

3.1. Design and synthesis of conjugates

The sequences of conjugates **1**–**4** were designed based on the target site of Chb-M′ (chlorambucil-β-IPII-γ-PPPP-Dp)[10], 5′-WCCWCW-3′, and the structure of conjugates **1** and **3** mimicked AcIIPP-γ-II-indole-*seco*-CBI.[21] For the Bind-n-Seq analysis, biotinylated *seco*-CBI-PI polyamide conjugates **1** and **2** were synthesized (Scheme 1). 2-Cl-Trt-Cl resin was used for P-loading.[31] After completion of couplings, the resin was treated with hexafluoroisopropanol for cleavage and indole-*seco*-CBI was coupled. As the pyrrole at the C-terminus experienced decarboxylation faster than deprotection when treated with 20% TFA in CH_2Cl_2, indole-*seco*-CBI was coupled, then the Boc group was removed by TFA. The last step was biotinylation to yield conjugates **1** and **2**. Conjugates **3** and **4** for cytotoxicity assays were prepared using a similar procedure (Scheme 1).



Scheme 1. Synthesis of conjugates **1**–**4**. Reagents and conditions: (i) 30% hexafluoroisopropanol, CH_2Cl_2; (ii) indole-*seco*-CBI, PyBOP, DIEA, DMF; (iii) 50% TFA, CH_2Cl_2; (iv) NHS-PEG_{12}-biotin, DIEA, DMF.

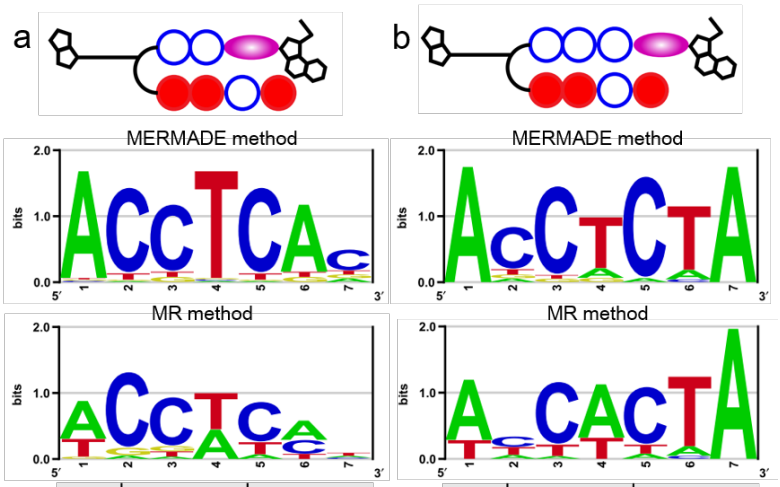
3.2. Bind-n-Seq analysis

3.2.1. MERMADE method

The sliding window for identifying binding motifs was set at 7-mer for both conjugates for comparison. Conjugates **1** and **2** showed motif logos successfully following the pairing rules[4-7] and the expected adenine at the alkylating site (Figure 2): 5′-WCCWCAN-3′ and 5′-WCCWCWA-3′, respectively (W = A or T; N = A, T, G or C). As the binding motif of conjugate **1** is only 6-mer, the seventh base was out of that 6-mer and is reasonably CTGA. The sixth alkylation base of conjugate **1** was not exclusively A, while the alkylation bases of conjugate **2** are exclusive. This finding may suggest the lower alkylating activity of conjugate **1**. The enrichment values are calculated as ratios of sample vs input control, but the reason for low enrichment values for **1** is unknown.

3.2.2. MR method[24]

This analytical method was developed in 2016 to compensate for weakness of the MERMADE method. On one hand as discussed in that paper, the MERMADE method has a tendency to overestimate the first-ranked sequence because it takes sequences with only one mismatch compared from that central sequence into consideration, resulting in unexpected selectivity of A or T. On the other hand, as the MR method picks up sequences from top to a designated rank, the representation becomes fairer to show A and T with comparable sizes. After a slight modification of the reported MR program in order to detect sliding of sequences from their reference more accurately, this feature was observed in this study as well (Figure 2). Still, it is noteworthy that the seventh base in motif of conjugate **2** was exclusive A, suggesting that this base was alkylated by *seco*-CBI. This is the first time to apply the MR method to analysis of alkylating PI polyamide conjugates.



rank	sequence	enrichment
1	ACCTCAC	2.693
2	CCTCACC	2.655
3	TACCTCA	2.476
4	CACCTCA	2.443
5	CCACCTC	2.405
6	AGGTGAG	2.239
7	GAGGTTA	2.211
8	ACCTCAT	2.171
9	CTACCTC	2.116
10	ACCTCAA	2.072
11	GAGGTGA	2.057
12	GGTGAGA	2.057
13	CTCCACC	2.055
14	ACCTCAG	2.053
15	CCACTCC	2.043
16	ATACCTC	2.018
17	GTGAGGA	2.018
18	ACCTCTA	2.013
19	CGACCTC	2.004
20	ACTCCTC	2.004

rank	sequence	enrichment
1	ACCTCTA	6.961
2	ACCACTA	5.106
3	ATCACTA	4.942
4	CCTCTAG	4.453
5	TAGTGAA	4.336
6	ACCTCAA	4.192
7	CCACTAG	4.129
8	ACTAGTG	3.823
9	CTAGTGA	3.823
10	CCTCTAA	3.786
11	CTCTAGA	3.681
12	CACTAGA	3.641
13	TACACTA	3.634
14	TCACTAA	3.613
15	ATCTCTA	3.461
16	AGTGGTA	3.367
17	CCTCTAC	3.331
18	ACCATTA	3.282
19	CCTCAAG	3.193
20	CCACTAA	3.163

Figure 2. Bind-n-Seq results obtained with conjugates **1** (a) and **2** (b). The top panel shows schematic representations of the conjugates, the middle panel motif logos by the MERMADE or MR method, and the bottom panel, tables of sequences numbered in order of enrichment values. The results by the MR method were obtained from top 20 excluding slid sequences, and the reference sequences were 5'-WCCWCAN-3' and 5'-WCCWCWCA-3' for **1** and **2**, respectively.

3.3. Cytotoxicity

To explore the utility of these *seco*-CBI-PI polyamide conjugates, we evaluated the cytotoxicity of conjugates **3** and **4** against A549 cells, a human lung carcinoma cell line, using the WST assay, and determined the IC_{50} values (the concentration required for 50% inhibition of cell-growth) of each conjugate (Figure 3). Conjugate **4** showed higher cytotoxicity than conjugate **3**. Nevertheless, both of the PIP-CBI conjugates showed IC_{50}

values at submicromolar concentrations. Chb-M′ targeting the same sequence as conjugates **3** and **4** was cytotoxic for A549 cells at micromolar concentrations.[10] These results suggest that *seco*-CBI-PI polyamide conjugates had higher cytotoxicity than the chlorambucil-PI polyamide conjugate.

Conjugate	**3**	**4**
IC_{50} /nM	120	63

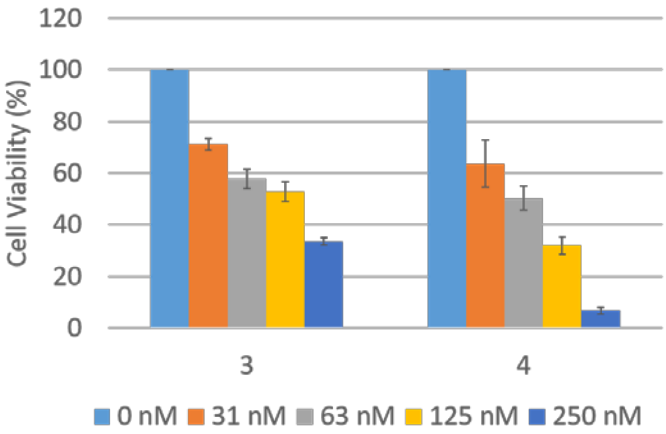


Figure 3. IC_{50} values of *seco*-CBI-PI polyamide conjugates against A549 cells.

4. Conclusions

As RUNX-targeting Chb-M′ enabled leukemia mouse models to survive much longer than the DMSO control, chlorambucil, the anticancer drug cytarabine, and PI polyamide control, CBI-version of PI polyamide conjugates were designed, synthesized, and characterized in vitro. Bind-n-Seq analysis with both the MERMADE and MR methods indicated that the two conjugates **1** and **2** had high sequence-specificity and alkylating activity. IC_{50} values for the A549 cancer cell line of around 100 nM were lower than those for Chb-M′. This study adds insight into the potential use of PI polyamide conjugates for cancer therapy.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number JP16H06356, Basic Science and Platform Technology Program for Innovative Biological Medicine by Japan Agency for Medical Research and Development (AMED), and the Platform Project for Supporting Drug Discovery and Life Science Research funded by AMED.

Supporting information

Supporting information for this article can be found under http://XXX.

References

1. Medscape from “http://reference.medscape.com/drugs/antineoplastics-alkylating”

2. Hurley LH. *Nat Rev Cancer.* 2002;2;188.

3. Yoon JH, Lee CS. *Arch Pharm Res.* 1997;20;550.

4. Trauger JW, Baird EE, Dervan PB. *Nature.* 1996;382;559.

5. Dervan PB. *Bioorg Med Chem.* 2001;9;2215.

6. Dervan PB, Edelson BS. *Curr Opin Struct Biol.* 2003;13;284.

7. Blackledge MS, Melander C. *Bioorg Med Chem.* 2013;21;6101.

8. Wang YD, Dziegielewski J, Wurtz NR, Dziegielewska B, Dervan PB, Beerman TA. *Nucleic Acids Res.* 2003;31;1208.

9. Dickinson LA, Burnett R, Melander C, Edelson BS, Arora PS, Dervan PB, Gottesfeld JM. *Chem Biol.* 2004;11;1583.

10. Morita K, Suzuki K, Maeda S, Matsuo A, Mitsuda Y, Tokushige C, Kashiwazaki G, Taniguchi J, Maeda R, Noura M, Hirata M, Kataoka T, Yano A, Yamada Y, Kiyose H, Tokumasu M, Matsuo H, Tanaka S, Okuno Y, Muto M, Naka K, Ito K, Kitamura T, Kaneda Y, Liu PP, Bando T, Adachi S, Sugiyama H, Kamikubo Y. *J Clin Invest.* 2017, doi: 10.1172/JCI91788.

11. Ito Y, Bae SC. Chuang LSH. *Nat Rev Cancer.* 2015;15;81.

12. Boger DL, McKie JA. *J Org Chem.* 1995;60;1271.

13. Boger DL, Ishizaki T, Kitos PA, Suntornwat O. *J Org Chem.* 1990;55;5823.

14. Bando T, Sugiyama H. *Acc Chem Res.* 2006;39;935.

15. Bando T, Sasaki S, Minoshima M, Dohno C, Shinohara K, Narita A, Sugiyama H. *Bioconj Chem.* 2006;17;715.

16. Minoshima M, Bando T, Shinohara K, Kashiwazaki G, Nishijima S, Sugiyama H. *Bioorg Med Chem.* 2010;18;1236.

17. Hiraoka K, Inoue T, Taylor RD, Watanabe T, Koshikawa N, Yoda H, Shinohara K, Takatori A, Sugimoto H, Maru Y, Denda T, Fujiwara K, Balmain A, Ozaki T, Bando T, Sugiyama H, Nagase H. *Nat Commun.* 2015;6;6706.

18. Lin J, Hiraoka K, Watanabe T, Kuo T, Shinozaki Y, Takatori A, Koshikawa N, Chandran A, Otsuki J, Sugiyama H, Horton P, Nagase H. *PLoS ONE.* 2016;11;e0165581.

19. Taylor RD, Chandran A, Kashiwazaki G, Hashiya K, Bando T, Nagase H, Sugiyama H. *Chem Eur J.* 2015;21;14996.

20. Anandhakumar C, Kizaki S, Bando T, Pandian GN, Sugiyama H. *ChemBioChem.* 2015;16;20.

21. Chandran A, Syed J, Taylor RD, Kashiwazaki G, Sato S, Hashiya K, Bando T, Sugiyama

H. *Nucleic Acids Res.* 2016;44;4014.

22. Zykovich A, Korf I, Segal DJ. *Nucleic Acids Res.* 2009;37;e151.

23. Sawatani Y, Kashiwazaki G, Chandran A, Asamitsu S, Guo C, Sato S, Hashiya K, Bando T, Sugiyama H. *Bioorg Med Chem.* 2016;24;3603.

24. Kashiwazaki G, Chandran A, Asamitsu S, Kawase T, Kawamoto Y, Sawatani Y, Hashiya K, Bando T, Sugiyama H. *ChemBioChem.* 2016;17;1752.

25. Meier JL, Yu AS, Korf I, Segal DJ, Dervan PB. *J Am Chem Soc.* 2012;134;17814.

26. Rodriguez R, Miller KM. *Nat Rev Genet.* 2014;15;783.

27. Kawamoto Y, Sasaki A, Chandran A, Hashiya K, Ide S, Bando T, Maeshima K, Sugiyama H. *J Am Chem Soc.* 2016;138;14100.

28. Chandran A, Syed J, Li Y, Sato S, Bando T, Sugiyama H. *ChemBioChem* 2016;17;1905.

29. Asamitsu S, Kawamoto Y, Hashiya F, Hashiya K, Yamamoto M, Kizaki S, Bando T, Sugiyama H. *Bioorg Med Chem.* 2014;22;4646.

30. Kawamoto Y, Sasaki A, Hashiya K, Ide S, Bando T, Maeshima K, Sugiyama H. *Chem Sci.* 2015;6;2307.

31. Li, BC, Montgomery DC, Puckett JW, Dervan PB. *J Org Chem.* 2013;78;124.

32. Minoshima M, Bando T, Sasaki S, Shinohara K, Shimizu T, Fujimoto J, Sugiyama H. *J Am Chem Soc.* 2007;129;5384.

33. Kashiwazaki G, Bando T, Yoshidome T, Masui S, Takagaki T, Hashiya K, Pandian GN, Yasuoka J, Akiyoshi K, Sugiyama H. *J Med Chem.* 2012;55;2057.

34. Schmieder R, Edwards, R. *Bioinformatics.* 2011;27;863.

35. Workman CT, Yin Y, Corcoran DL, Ideker T, Stormo GD, Benos PV. *Nucleic Acids Res.* 2005;33;W389.